Oracle Exadata: a Data Management Tipping Point

Merv Adrian, Principal, IT Market Strategy
www.itmarketstrategy.com

Architectural Innovation Delivers the Full Stack Solution

Data management technology is on the verge of a new era – and users who commit to its architectural innovations earlier will gain disproportionate rewards in performance, scalability, flexibility and cost. Advances and new directions in the components of data-management architecture – CPU, memory, storage, I/O, and of course the database engine itself – have been steady. But they have hitherto occurred separately, and when they are instead integrated effectively, the resulting synergy delivers a major step forward in the raw power of business- and mission-critical applications. The emerging next-generation platform will enable users to keep pace with the ever-increasing demand for more storage (data growth is literally outpacing the growth in available storage), faster response times, and growing user populations. Most important, the new platform will inspire and enable new, richer data-driven applications in areas such as Enterprise Performance Management (EPM).

Integrating the new capabilities is not straightforward. For example, solid-state devices (SSDs) are increasingly useful as an additional storage tier between main memory and disk that can speed delivery of data to the CPU. However, unless the other parts of data management architecture are designed to optimize a storage hierarchy with SSDs in it, many of the performance benefits of SSDs are lost. A holistic approach to design must be in place to gain all the benefits of advances in every part of the architecture.

This White Paper describes Oracle's Exadata V2 solution, also known as the Sun Oracle Database Machine. In this second version of its newest platform, Oracle not only provides the latest technology in each part of the data-management architecture, but also integrates them under the full control of one vendor, with a unified approach to leveraging the full stack. Early indications are that Oracle Exadata customers are already seeing the benefits of this approach.

Copyright March 2010, Merv Adrian, IT Market Strategy

Greater Than the Sum of its Parts

Breakthrough advances in data management have been enabled by technology innovations in CPUs, memory, storage, and I/O. Below, we consider each separately. Individually, these developments enable new approaches to performance optimization; together, they empower a step function uplift. Database software today is stuffed with features designed to surmount the obstacles imposed by earlier limits on processor speed, memory availability, and especially the movement of data from physical storage into the layer of memory closest to the processor (I/O). Software designed to exploit the newest architectures can take innovative approaches to processing, unconstrained by many of these limitations, and improve performance. Taken together, the architectural transformations enable much more: they redraw the map of possibilities for both business intelligence and transaction processing.

CPU

Recent technology advances that empower data-management performance boosts include multicore chips, far more extensive use of prefetch and pipelining, and 28-nanometer chip-fabrication processes that increase per-core speed. Multicore and pipelined CPU architectures enable much greater parallelization of transactions. Oracle Exadata V2 uses a pair of the latest Intel Xeon quad-core chips in every server, whether it be a database server or the Oracle Exadata Storage Server, to achieve 8-core parallelization, with prefetch and pipelining. This, in turn, can produce a two-times-or-more improvement in the time to process, say, a 1-terabyte query.

Memory

Most of the improvements in this area have to do with cost, memory speed and capacity. Over the last few years these advances have meant gigabyte- and even terabyte-sized data-warehouse memory sizes, and more extensive use of in-memory database technology have become commercially feasible. One benefit is a reduced dependency on disk lookup for data which is read frequently but rarely updated – such as lookup tables. This approach reduces processing time and eliminates bottlenecks. The Oracle Exadata Database Server Grid offers 72 gigabytes of memory in each database server, for a total of 576 gigabytes of memory in an 8-server, full-rack configuration. Leveraging this enormous memory space effectively requires more than simply "dropping the tables into memory" – even the basic steps used by the cost-based optimizer will benefit, and large parts of the code path can be eliminated in the earliest processing stages.
For example, Oracle Database 11g Release 2 automatically distributes a table across buffer caches for in-memory parallel queries if the optimizer determines frequent access would thus be improved.

I/O

I/O is a perpetual issue, and the recent explosion of data has compounded its challenges. Using flash memory for read caching can eliminate a great deal of I/O completely, and Oracle Exadata uses this strategy effectively as well as offering

Copyright March 2010, Merv Adrian, IT Market Strategy

both SAS and SATA attached storage, with a 40 Gb/s Infiniband interconnect with RDS/RDMA (remote direct memory access) to dramatically improve the movement of data from storage to the database. Leveraging Oracle Database 11g's Automatic Storage Management (ASM) feature, a purpose-built, vertically integrated file system and volume manager provides the performance of asynchronous I/O with the easy management of a file system to help users scale out gracefully. At the same time, CPU performance improvements noted above permit data-management applications to deliver much higher very-large-data-store performance by tackling the increased load gracefully. In addition to its raw performance benefits, this ability to move data more rapidly to where it is needed lays the groundwork for continuing innovation in multilevel storage strategies that keep less-used data in less expensive form factors and compress it more aggressively to reduce costs.

Storage

There has been steady improvement in storage density and transfer speed in recent years as well as gains from more sophisticated RAID striping for parallel access to groups of records, and online snapshotting of disks to take some of the data-store backup/recovery burden from the database engine. All of this innovation is driven by ever-more sophisticated (and complex) software, sometimes inside the database code line itself.

Additional features of the Oracle Exadata Storage Server deliver further performance kicks. One is Smart Scan: the storage system itself can perform some queries on subsets of the data via the iDB protocol. Thus, in many queries predicates can be "pushed down" to the storage level for preliminary evaluation, eliminating the retrieval of unnecessary data.

At a deeper level, Storage Indexes create 1 megabyte "smart storage cells;" the index contains maximum and minimum values for each commonly queried column. If the value requested is "out of range," that block of data does not need be read at all.

But by far the most dramatic advance in storage technology in recent years has been the increased cost-effectiveness of solid-state disk (SSD), the use of non-volatile memory itself as a storage medium. SSDs are seeing wider use as a storage tier between memory and disk. Users are reporting an order-of-magnitude performance boost from the SSD's faster read/write times. Moreover, the improvement comes without substantial new software requirements; in fact, much sophisticated branching logic and cost optimization can be bypassed entirely, reducing the number of instructions to be executed before results are obtained.

Each Oracle Exadata Storage Server includes 384 GB of "flash memory" SSDs, and a full rack of 14 storage servers has over 5Tb. Oracle Exadata Storage Server coordinates with the database engine to use striping and other techniques to improve performance. Incremental backups execute faster due to fine-grained block change tracking available only on Exadata storage.

Copyright March 2010, Merv Adrian, IT Market Strategy

Similarly, flash cache "in front of" disks maintains data in use and allows processing to bypass disk controller-driven retrievals. This increases the speed of querying on those subsets, since the data does not need to be loaded to memory, and removes some of the CPU load. The new caching strategies do not flush the cache automatically for backup or table scans, eliminating "reloads" that take additional processing time.

Database

Traditional row-oriented relational databases are optimized for writes and random reads; columnar databases excel in queries across a large number of records where the values in a small number of columns are accessed. Columnar databases also leverage the commonality of content in the columns to drive substantial compression. A hybrid database seeks to achieve most of the performance advantages of both relational and columnar. Recently published benchmarks suggest that in some circumstances, columnar and/or hybrid technology can achieve order-of-magnitude increases in performance in certain types of large-scale (multi-terabyte) data warehousing.

Oracle Exadata V2 Hybrid Columnar Compression introduces a "disk-block" hybrid approach that allows users to specify that particular "compression blocks" of data are to be stored in columnar format simply by specifying hybrid columnar compression for a table or partition of a table. These blocks take advantage of the full spectrum of columnar technology: indexing, compression, and delayed expansion of the data until a result is about to be delivered. Moreover, the Oracle Exadata Storage Server can also apply more dense columnar compression techniques to archived data, ensuring better backup/recovery performance (less data to dump/load) and less frequent need to add storage.

Smart Scan can execute SQL actions such as SELECT and PROJECT at the storage level. So, for example, the system will read compressed data, extract only the needed columns, select on, say, the DATE column, then send results on without that column for SUM execution.

Copyright March 2010, Merv Adrian, IT Market Strategy

Delivering On the Promise: Cross-Stack Benefits

The real "secret sauce" of Oracle Exadata V2 is the way in which these technologies complement each other to deliver additional performance and scalability. Traditionally, server, storage, and database technologies from different vendors achieved excellent optimization within their own spheres, but that optimization often worked at cross-purposes with other parts of the data-management architecture. For example, database and storage vendors have traditionally optimized data storage for different user needs, so that a storage area network (SAN) solution aimed at remote-PC document and video data could undercut the optimization of a data-warehouse data store.

This integration shows in several areas. In querying, the database engine works in concert with the storage solution, with server and storage being allocated query tasks in a version of load balancing across CPU, SSD, and disk-array processing. Storage allocation and compression are integrated between the storage manager and the database engine. The CPU can operate on less data, as retrieval is assisted by pushdown processing in the storage layer.

Another key aspect of integration is flexibility in workload management: the ability to apply different data-management technologies seamlessly to different workloads within one overall solution. For example, while complex-query data-analysis solutions use hybrid-columnar technology up and down the integrated stack to deliver deeper analysis on smaller, compressed data sets that have been rapidly pre-filtered by smart storage, Oracle Exadata's OLTP write-heavy solutions can take advantage of the new SSDs to process transactions faster, with fewer buffer flushes, and less forced swapping to disk. These advances are transparent to applications and programmers; they require no application rewrite whatsoever.

Finally, the one-vendor, integrated approach allows "black-box" rapid deployment of pre-tested full solutions using Sun Oracle Database Machine and Exadata Storage Servers as building blocks in a grid architecture, with modular upgrade of storage and server together, from a quarter rack to a half, or a half to a full, based on need.

Application Potential

For many shops, the potential of Oracle Exadata lies in its ability to do what is already being done, but faster and cheaper, or in the assurance that they may now be able to keep ahead of ever-growing demands for more data. But an equally significant – and perhaps more exciting - use of Oracle Exadata is the promise of using its raw power to do something that hasn't been done before. One example of this might be Enterprise Performance Management (EPM).

EPM is an outgrowth of the old manual planning/budgeting/accounting cycle of corporate finance. Twenty years ago, it might take a month to do a yearly plan for a large-scale enterprise, and a month to record the quarterly "actuals" and come out with a quarterly set of balance sheet, income statement, and cash analysis.

Copyright March 2010, Merv Adrian, IT Market Strategy

Today, consolidation of weekly results from a 150-line-of-business global enterprise, followed by comparison with plan, replanning, and rebudgeting, can be done in less than a working day. Solutions such as Oracle Exadata V2 hold the potential of reducing this entire process to hours.

This speedup means that enterprise-wide plans are not cast in concrete for a full year; they can be modified, drastically if necessary, on a monthly or even weekly basis, when the organization changes strategy in midstream. But the real new EPM application potential is in using deeper what-if analysis during the consolidation process to drive more granular plan/budget modifications. In fact, the new EPM can even empower lines of business to carry out their own modifications based on their own what-if analyses on a weekly basis, within the corporate guidelines, while corporate senior finance personnel are generating overall results. In other words, solutions such as Oracle Exadata V2 allow EPM to move surprisingly close to real-time reaction to events.

The Customer Experience

While Oracle Exadata V2 is not in widespread use yet, initial customer reaction has been positive. For example, Banca Transilvania, a Romanian financial institution providing services for small to medium-scale enterprises, began moving its 5-terabyte data warehouse and reporting applications to V2 starting in December of 2009. Early testing has suggested that performance is about 10 times better than even Exadata V1 (and, in some cases, 30-50 times better than the previous solution), and that it cost only 20-25% as much to achieve the needed performance as with another data-management solution using other hardware and storage vendors. Indications are that these performance and price-performance advantages are translating into an ability to do deeper analysis and storage-upgrade savings from smaller data stores.

The customer reports that the upgrade to Exadata and to V2 has been going well, with migration of all targeted apps projected to take a total of about four months. Oracle has been "very responsive", and the bank anticipates adding Oracle Database Vault and Oracle Transparent Data Encryption capabilities in the near future.

TUI Netherlands, an integrated travel operator with 150 travel brands and hundreds of websites, implemented an Exadata V1 machine in December of 2009 for both data warehousing and OLTP, moving 4 terabytes of data in one overnight load with Oracle's help. In March 2010, TUI installed Exadata V2, moving the data in less than 3 hours (so far without data compression.) With the new configuration, TUI's reservation system (the OLTP piece) runs in memory and is dramatically faster. Where TUI recalculated the millions of price scenarios across hotel, travel and meal reservations "2-3 times per year on the old platform, we can now do it much more frequently" because of the performance improvements, responding more rapidly to changing business conditions.

Copyright March 2010, Merv Adrian, IT Market Strategy

Conclusions

In Oracle Exadata V2, Oracle has delivered a purpose-built machine for a wide variety of enterprise needs: data warehousing, OLTP and mixed workloads. Exadata makes extensive use of database innovations: Smart Scan, Smart Flash Cache, and Exadata Hybrid Columnar Compression combine to rewrite the rules of data movement and I/O within a large data environment. In the data center, Storage Indexes, Infiniband with RDS/RDMA and scale-out storage (ASM) leverage hardware advances to drive new ways to leverage and grow the use of existing applications without reprogramming or re-architecting. Collectively, these innovations afford users new potential, for example by supporting deeper analysis for data mining, and in evolving applications such as EPM to achieve new real-time strategy-change or rapid-reaction benefits.

Early customer results indicate that Oracle Exadata V2 is delivering on the promise of those new capabilities. The action item for IT, then, is to target corporate applications with the biggest potential for speed-based enhancement, and consider whether Oracle Exadata V2 can cost-effectively enable new ways to use these for competitive advantage.

Copyright March 2010, Merv Adrian, IT Market Strategy